Insights into the Origin of Atlantic Salmon Juveniles Observed in the Mispec River

P.T. O'Reilly and M. Cassista-Da Ros

Department of Fisheries and Oceans
Science Branch, Population Ecology Division
Bedford Institute of Oceanography
1 Challenger Drive, P.O. Box 1006
Dartmouth Nova Scotia
B2Y 4A2

2009

Canadian Technical Report of
Fisheries and Aquatic Sciences 2856

Fisheries and Oceans Canada | Pêches et Océans Canada

Canada

Canadian Technical Report of Fisheries and Aquatic Sciences

Technical reports contain scientific and technical information that contributes to existing knowledge but which is not normally appropriate for primary literature. Technical reports are directed primarily toward a worldwide audience and have an international distribution. No restriction is placed on subject matter and the series reflects the broad interests and policies of Fisheries and Oceans Canada, namely, fisheries and aquatic sciences.

Technical reports may be cited as full publications. The correct citation appears above the abstract of each report. Each report is abstracted in the data base *Aquatic Sciences and Fisheries Abstracts*.

Technical reports are produced regionally but are numbered nationally. Requests for individual reports will be filled by the issuing establishment listed on the front cover and title page.

Numbers 1-456 in this series were issued as Technical Reports of the Fisheries Research Board of Canada. Numbers 457-714 were issued as Department of the Environment, Fisheries and Marine Service, Research and Development Directorate Technical Reports. Numbers 715-924 were issued as Department of Fisheries and Environment, Fisheries and Marine Service Technical Reports. The current series name was changed with report number 925.

Rapport technique canadien des sciences halieutiques et aquatiques

Les rapports techniques contiennent des renseignements scientifiques et techniques qui constituent une contribution aux connaissances actuelles, mais qui ne sont pas normalement appropriés pour la publication dans un journal scientifique. Les rapports techniques sont destinés essentiellement à un public international et ils sont distribués à cet échelon. Il n'y a aucune restriction quant au sujet; de fait, la série reflète la vaste gamme des intérêts et des politiques de Pêches et Océans Canada, c'est-à-dire les sciences halieutiques et aquatiques.

Les rapports techniques peuvent être cités comme des publications à part entière. Le titre exact figure au-dessus du résumé de chaque rapport. Les rapports techniques sont résumés dans la base de données *Résumés des sciences aquatiques et halieutiques*.

Les rapports techniques sont produits à l'échelon régional, mais numérotés à l'échelon national. Les demandes de rapports seront satisfaites par l'établissement auteur dont le nom figure sur la couverture et la page du titre.

Les numéros 1 à 456 de cette série ont été publiés à titre de Rapports techniques de l'Office des recherches sur les pêcheries du Canada. Les numéros 457 à 714 sont parus à titre de Rapports techniques de la Direction générale de la recherche et du développement, Service des pêches et de la mer, ministère de l'Environnement. Les numéros 715 à 924 ont été publiés à titre de Rapports techniques du Service des pêches et de la mer, ministère des Pêches et de l'Environnement. Le nom actuel de la série a été établi lors de la parution du numéro 925.

Canadian Technical Report of
Fisheries and Aquatic Sciences 2856

2009

INSIGHTS INTO THE ORIGIN OF ATLANTIC SALMON JUVENILES OBSERVED IN THE MISPEC RIVER

by

P. T. O'Reilly and M. Cassista-Da Ros

Science Branch
Fisheries and Oceans Canada
Population Ecology Division
Bedford Institute of Oceanography
P.O. Box 1006, 1 Challenger Drive
Dartmouth, NS B2Y 4A2
E-mail: Patrick.Oreilly@dfo-mpo.gc.ca

© Her Majesty the Queen in Right of Canada, 2009.
Cat. No. Fs 97-6/2856E ISSN 0706-6457
PDF Version: Cat. No. Fs97-6/2856-PDF ISSN 1488-5379

Correct citation for this publication:

O'Reilly, P.T., and M. Cassista-Da Ros. 2009. Insights into the Origin of Atlantic Salmon Juveniles Observed in the Mispec River. Can. Tech. Rep. Fish. Aquat. Sci. 2856: iv + 15 p.

TABLE OF CONTENTS

Abstract / Resumé iv

Introduction 1

Materials and Methods 3

Results 5

Discussion 7

Literature Cited 8

LIST OF TABLES

Table 1. Sample collection information for 19 Atlantic salmon juveniles obtained from the Mispec River and submitted to the Department of Fisheries and Oceans for genetic analyses, including length, weight, and the general and specific location at which samples were obtained 10

Table 2. Marker type employed (nuclear versus mitochondrial), mode of inheritance, number of alleles observed, and locus references providing primer sequences and annealing temperatures for microsatellite loci 11

Table 3. Microsatellite and mitochondiral DNA (mtDNA) Single Nucleotide Polymorphism (SNP) data, and possible first-order relatedness as inferred from Kinship analyses, for 17 Atlantic salmon samples analyzed 12

Table 4. Distribution of 17 multilocus genotypes into full- and half-sib groupings for each of 10 simulated datasets 13

Table 5. Likelihood of detecting inner Bay of Fundy mitochondrial DNA types given variable population allele frequencies and different effective numbers of juveniles sampled 14

LIST OF FIGURES

Figure 1. Geographic locations of the Mispec and other reference rivers in the Bay of Fundy area 15

ABSTRACT

O'Reilly, P.T., and M. Cassista-Da Ros. 2009. Insights into the Origin of Atlantic Salmon Juveniles Observed in the Mispec River. Can. Tech. Rep. Fish. Aquat. Sci. 2856: iv + 15 p.

The number of Atlantic salmon adults returning to rivers of the inner Bay of Fundy has declined precipitously in recent years, to the point that juvenile salmon are rarely detected in rivers that are not supported by captive breeding and rearing efforts. When juveniles are observed in non-supported rivers, there is often considerable interest as to the origins of these individuals, both because of possible implications for the restoration of salmon to the inner Bay, but also because of potential legal ramifications of findings of endangered salmon in rivers identified as potential habitat for this group of populations. In 2007, 19 juveniles were observed and sampled from the Mispec River. In 2008, microsatellite and mitochondrial DNA Single Nucleotide Polymorphism (SNP) analyses were carried out on 17 of these 19 samples. No mitochondrial DNA haplotypes specific to, and diagnostic of, the inner Bay of Fundy assemblage of Atlantic salmon were observed in this collection. Kinship and population analyses indicate that the set of juveniles observed was likely produced by a large number of females (11 to 17 and most probably around 13), indicating modest (0.77) to (0.99) high (0.99) power to detect this mitochondrial DNA haplotype in the population from which these juveniles originated. The presence of certain microsatellite alleles at the microsatellite locus *Ssa*202 not observed in large population surveys of the inner Bay of Fundy carried out previously further suggests that these individuals are not progeny of Atlantic Salmon from the inner Bay of Fundy, but rather are probably offspring of either strays from the nearby Saint John River system, or escapes from the aquaculture industry.

RÉSUMÉ

O'Reilly, P.T., and M. Cassista-Da Ros. 2009. Insights into the Origin of Atlantic Salmon Juveniles Observed in the Mispec River. Can. Tech. Rep. Fish. Aquat. Sci. 2856: iv + 15 p.

Le nombre de saumons atlantiques qui retournent aux rivières de la baie intérieure de Fundy a radicalement diminué ces dernières années, à tel point qu'il est maintenant rare de trouver des saumons juvéniles dans les rivières qui ne bénéficient pas du soutien des programmes de sélection et d'élevage en captivité. Quand des saumons juvéniles sont observés dans des rivières qui ne bénéficient pas de ces programmes, leur origine suscite souvent énormément d'intérêt, en raison, d'une part, de ce que peut signifier leur présence pour le rétablissement du saumon dans la baie intérieure de Fundy et, d'autre part, des répercussions juridiques possibles de la présence de saumons de populations en voie de disparition dans des rivières considérées comme un habitat possible de ce groupe de populations. En 2007, 19 juvéniles ont été observés dans la rivière Mispec et des échantillons de leurs tissus ont été prélevés. En 2008, on a procédé à des analyses de microsatellites et d'ADN mitochondrial par étude de SNP (polymorphismes de nucléotide simple) sur 17 de ces 19 échantillons. Ces analyses n'ont pas décelé d'haplotypes d'ADN mitochondrial propres à l'assemblage de saumons atlantiques de la baie intérieure de Fundy et révélateurs de la présence de tels saumons. Les analyses de population et de liens de parenté qui ont été effectuées avec les juvéniles observés étaient vraisemblablement issus de nombreuses femelles (entre 11 et 17 et très probablement aux environs de 13), ce qui se traduit par une puissance statistique allant de modeste (0,77) à élevée (0,99) pour déceler cet haplotype d'ADN mitochondrial dans la population d'où proviennent ces juvéniles. La présence de certains allèles de microsatellite au locus microsatellite *Ssa*202, qui n'avait pas été observée dans de vastes relevés antérieurs sur les populations de la baie intérieure de Fundy, nous suggèrent que les individus observés ne sont pas la progéniture du saumon atlantique de la baie intérieure de Fundy, mais plutôt des descendants soit de saumons égarés provenant du bassin versant de la rivière Saint-Jean, situé à proximité, soit de saumons évadés de sites aquacoles.

INTRODUCTION

Atlantic salmon inhabiting the 40 or more rivers and streams to the east of New Brunswick's Saint John and Nova Scotia's Annapolis rivers exhibit unique life history traits, including high incidence of maturation after only one sea-winter and a high incidence of repeat spawning (Ducharme, 1969), and possible local migration as suggested by the recovery of tags from geographically proximate locations in the Bay of Fundy and Gulf of Maine (Jessop, 1976). Whereas approximately 40,000 adult salmon once returned to rivers and streams of the inner Bay (iBoF) to spawn each fall, fewer than an estimated 250 returned in 2000 (DFO, 2006). In fact, electrofishing surveys carried out in 2000 and 2002 failed to detect any fry in 30 of 34 rivers surveyed that were not supported by captive breeding and rearing (Gibson *et al.*, 2003). Possible origins of juvenile salmon in the four non-supported rivers where salmon were detected include 1) persistence of small residual native populations not detected in previous surveys, 2) straying of iBoF salmon from nearby rivers and successful spawning one or two years previously, 3) straying of salmon from the still large Saint John population and successful spawning one or two years previously, 4) straying of wild salmon from more distant sources and successful spawning one or two years previously, 5) straying of nearby aquaculture salmon and successful spawning one or two years previously, and 6) the unauthorized release of juveniles from aquaculture facilities nearby.

Atlantic salmon juveniles have recently been captured in the Mispec River, a drainage that has supported salmon historically, and which is within the geographic boundaries associated with inner Bay of Fundy salmon (DFO, 2006). However, the Mispec River is very much a peripheral inner Bay river, located just 10 to 15 kilometres to the east of the large Saint John River, an outer Bay of Fundy drainage (Figure 1). The Mispec is also geographically proximate to the Passamoquoddy and Cobscook Bay areas, sites of large-scale rearing of farmed Atlantic salmon in marine net pens (DFO, 1999). In order to investigate whether juveniles observed in the Mispec River are progeny of true iBoF salmsn versus offspring of either natural strays from the outer Bay of Fundy or aquaculture escapees from nearby marine sea pens, tissue samples were collected from 17 Atlantic salmon juveniles and sent to the Department of Fisheries and Oceans for microsatellite and mitochondrial DNA (mtDNA) analyses.

When molecular genetic differences among populations are very large, with some variants fixed in one population but absent (or nearly so) in others, mere presence/absence information on particular genetic variants can be used to infer population or subspecies of origin. Typically, however, genetic differences are less marked, with populations exhibiting somewhat overlapping sets of alleles, varying only in terms of relative proportions of each. In such instances, allele frequency information from multiple genetic markers (typically 10-15), and likelihood-based assignment-test methods, are used to estimate population of origin of unknown samples (recently reviewed in Manel *et al.*, 2005). However, even here, existing allele frequency differences of sufficient magnitude must exist between putative populations of origin, and allele frequency information must be available from all potential sources or populations, to accurately assign individuals to populations.

The situation for salmon of the Bay of Fundy area is somewhat unique, in that populations exhibit *very* little differentiation at nuclear microsatellite markers surveyed to date, with between-population F_{ST} values less than 0.05 (unpublished data), a level of structuring at which assignments of individuals is untested and problematic (Manel *et al.*, 2005 and references

within). Yet, iBoF salmon often exhibit a mitochondrial DNA (mtDNA) haplotype, henceforth referred to as the iBoF mtDNA haplotype, not observed in several dozen salmon analyzed from the Saint John river system (Verspoor *et al.*, 2002). That this mtDNA haplotype is indeed very rare or absent elsewhere along the east coast of North America, is indicated by it's absence in large-scale surveys by Verspoor *et al.* (2005), and King *et al.* (2000), collectively involving thousands of salmon from dozens of rivers in multiple regions. Additionally, Single Nucleotide Polymorphim (SNP) analyses of salmon from several tributaries of the neighbouring Saint John River system failed to detect even one "T" base at site 4199 in 83 individuals surveyed (O'Reilly, unpublished data); the inner Bay mtDNA haplotype exhibits a "T" at base 4199 in the ND1 gene (see Hurst *et al.*, 1999 for DNA sequence and base position information, and Verspoor *et al.*, 2002 for complete haplotype information).

Although it seems likely that the iBoF mtDNA haplotype may not exist outside of the inner Bay, the iBoF mtDNA haplotype is not an ideal marker for identifying iBoF salmon in mixed or unknown sample collections because not all native wild salmon sampled from inner Bay rivers exhibit the iBoF mtDNA type identified by Verspoor *et al.* (2002). In the limited analysis conducted by Verspoor *et al.* (2002), allele frequencies were observed to vary between 0 and 75% among samples analyzed from different iBoF rivers. In fact, the iBoF mtDNA type was not detected in any of the relatively few samples obtained from three Chignecto Bay rivers analyzed in their original study (Verspoor *et al.*, 2002), though in a later study, the iBoF haplotype was observed in 3 of 7 fish analyzed from the Gardeners River, a Chignecto Bay drainage located between the Mispec and Irish rivers (Verspoor *et al.*, 2005). Furthermore, SNP analyses of dozens of juvenile salmon collected from the Point Wolfe and Big Salmon rivers have identified mtDNA exhibiting the base "T" at site 4199 in 27% and 11% of the salmon surveyed, respectively (O'Reilly, unpublished data).

Despite the fact that the iBoF mtDNA haplotype is not present in all iBoF salmon, analyses of variation at the ND1 4199 site may still provide insight into the origins of salmon observed in the Mispec River. If a "T" base at site 4199 is observed in at least one of the salmon submitted, this would indicate that at least one parent was not an aquaculture or wild Saint John salmon, but instead originated from an inner Bay river, either from a possible residual Mispec population or from another still extant inner Bay population, such as the Big Salmon or Stewiacke. Where sample collections include multiple individuals, the absence of iBoF mtDNA haplotypes can also be informative; under certain conditions, the absence of iBoF haplotypes can be taken as evidence for a non-inner Bay origin of the group of individuals sampled, though with some caveats. The likelihood of observing an iBoF mtDNA haplotype, had these fish originated from some inner Bay source, is a function of 1) the frequency of the mtDNA haplotype in the iBoF population from which these fish originated and 2) the number of individuals sampled. For example, if the frequency of the iBoF mtDNA type in a hypothetical population is 0.5, and 5 unrelated individuals are sampled, then the probability of not observing at least one mtDNA haplotype in the collection would be 0.5^5 or 0.03125. Since this is very unlikely, the absence of iBoF mtDNA, in this circumstance, could be taken as fairly strong evidence that the sample collection was not obtained from the hypothetical population in question.

An important caveat in this interpretation of the above results is the possible existence of family structuring in the sample collection obtained, of particular concern given the maternal mode of inheritance of mtDNA. For example, all offspring of a single female and a dozen or more males will exhibit the same single mtDNA type as that observed in the female parent. Thus, in the

example above, the likelihood of not observing an iBoF mtDNA haplotype would be 0.5^1 (0.5) and not 0.5^5 (0.03125).

To assess and mitigate the effects of possible family structuring on the interpretation of results, we estimated first-order relatedness in the collection of 17 Mispec juveniles using multilocus genotype information and the likelihood-based program *Colony* (Wang, 2004). Because *Colony* and other kinship reconstruction programs, such as *Pedigree* (Butler *et al.*, *2004)*, tend to artificially create some small full- or half-sib groupings (Wang, 2004; Smith *et al.*, 2001), we tested kinship reconstruction methods employed here on multiple sets of 17 simulated "unrelated" multilocus genotypes, drawn at random from the Mispec pool of sampled alleles, as in Hansen and Jensen (2005). These analyses provide an indication as to whether artificial or false groupings could be expected in a limited sample of 17 multilocus genotypes under conditions encountered here and, if so, insight into their nature and expected number.

In this analysis, in addition to reporting results on the absence/presence (and frequency) of the iBoF mtDNA haplotype in the Mispec samples submitted, we assess the power of these analyses to detect iBoF mtDNA haplotypes in the Mispec samples submitted given the implications of the above kinship analyses, but also under several mtDNA haplotype frequencies observed in different iBoF populations. This analysis was important in interpreting the implications of the possible finding of the absence of mtDNA in these samples. We also discuss observations of a particular microsatellite allele observed at the locus *Ssa*202, and implications of these findings on the origin of the juveniles observed in the Mispec River in 2007.

MATERIALS AND METHODS

Sample collection information: Nineteen juvenile Atlantic salmon were obtained from multiple sites from the Mispec River in September of 2007 (Table 1). For each individual collected, total length (millimetres), fork length (millimetres), and weight (grams) were recorded (Table 1), and a sample of caudal or adipose fin removed and stored in ethanol for later molecular genetic analyses.

Laboratory analyses: Fin clips were collected and stored in 1.5 ml microcentrifuge tubes containing 1 ml of ethanol. DNA was extracted and purified using Qiagen's 96-well DNeasy kits, following the manufacturer's specifications. Polymerase Chain Reaction (PCR) amplifications were carried out in 10 µl volumes, containing 1-100 nanograms of template DNA, 0.2 mM each dNTP, 0.1 µM labelled and unlabelled primers, 1X KCl buffer (10mM Tris HCl, 50 mM KCl, 0.08% Nonidet P40), 2 units of Taq DNA polymerase supplied by MBI Fermentas and 2.5 mM $MgCl_2$. Thermal cycling conditions were as follows: (94°C for 3 min)X1, (94°C for 45 sec, 58°C for 30 sec, 74°C for 1 min)X9, and (94°C for 30 sec, 58°C for 30 sec, and 72°C for 1 min)X 27, followed by a 30 min extension step at 72°C. Primer sequences for loci *Ssa*171, *Ssa*197 and *Ssa*202 are given in O'Reilly *et al.* (1996); *SSsp*2201, *SSsp*2210, *SSsp*2215, *SSsp*2216, *SSsp*1G7 and *SSsp*1605 are given in Paterson *et al.* (2004); and *Ssa*D58, *Ssa*D144 and *Ssa*D486 in King *et al.* (2005); additional information on individual loci are given in Table 2. Following amplification, PCR products were combined, and salt, unincorporated dNTPs, and unincorporated labelled and non-labelled primers were removed using Qiagen's 96-well MinElute plates, as specified by the manufacturer. Fragments were size-fractionated and detected using an Applied Biosystems 3130XL. Samples were cross-standardised between

platforms and between data sets by including two individuals from other reference collections. One sample from each strip of eight tubes was duplicated in wells 87 to 96 to identify sample placement errors, strip inversions, and plate inversions.

Single Nucleotide Polymorphism analyses of the ND1 4199 site were carried out using the *Invader* methodology from Third Wave technologies, Madison, WI (Lyamichev *et al.*, 1999). First, a segment of the mitochondrial ND1 locus spanning 4095-4276 base pairs (bp) (bp positions as given in Hurst *et al.*,1999) was amplified using the primers 5' TGT GCC ACT GCT CGT AGA 3' and 5' GCC CTT ACG CTT GCA CTA 3'. The amplification reagents are as described above for microsatellites with the exception of a buffer change to a 1X $(NH_4)_2SO_4$ solution (75 Tris HCl, 20 mM $(NH_4)_2SO_4$, 0.01% Tween 20), an increase in concentration of labelled and unlabelled primers to 0.5 µM, and an increase in the amount of Taq polymerase to 2.5 units per 10 µl PCR. Thermal cycling conditions were as follows: (95°C for 1 min)X1, (94°C for 1 min, 53°C for 1 min, 72°C for 1.5 min)X30, and a 5 min extension step at 72°C. The resulting PCR products were then diluted 1 in 50, and 7.5 µl combined with an equivalent volume of Invader reagent supplied by the manufacturer. The mixture was then incubated at 95°C for 3 minutes to denature DNA, then 10 minutes at 63°C for completion of the Invader reactions, then cooled to 10°C before reading fluorescence on a BMG Fluoristar Optima Fluorescence plate reader. The 15 µl volumes in the individual wells were excited at 485 and 544 nanometres, and emissions detected at 520 and 590 nanometres, respectively. Base calls were determined from relative fluorescence values in sample wells, and comparisons with non-template and template controls, and algorithms provided by Third Wave technologies.

Kinship analyses: First-order relatedness at the full-sib level (offspring sharing both a common male and female parent), and half-sib level (offspring sharing a common maternal or paternal parent), was assessed in 1) the group of 17 Atlantic salmon sampled from the Mispec and 2) 10 groups of simulated multilocus genotypes, using the program *Colony* (Wang, 2004). *Colony* uses a likelihood-based approach, allele frequency information from growing groups of siblings, and rules of Mendelian inheritance to cluster individuals. The Class I (upper allele dropout) and Class II (stochastic) error rates in all *Colony* analyses were set to 0.01. Additional details of kinship reconstruction methods used here are available in Herbinger *et al.* (2006).

Simulation analyses: Ten sets of 17 multilocus genotypes were generated by randomly sampling individual microsatellite alleles (with replacement) from the pool of microsatellite alleles obtained from the Mispec River, assuming Hardy Weinberg equilibrium and non- linkage of all microsatellite loci. The odd missing genotype in the original collection of samples was incorporated into the simulation analyses by creating an identical pattern of missing genotypes (one individual missing two specific single-locus genotypes, and three individuals missing one specific single-locus genotype). The simulated datasets were created using the program PopTools, developed by G. Hood of the Pest Animal Control Co-operative Research Center, Crisco, Canberra, Australia, and available at http://www.cse.csiro.au/poptools/index.htm. The degree to which the algorithm used in *Colony*, and the run parameters specified in the original analyses, falsely identified relatedness in the original true dataset was evaluated by testing for relatedness in these simulated sets of multilocus genotypes, using *Colony*.

Power analysis and the ability to detect an iBoF mtDNA haplotype in the Mispec sample collection: Theoretically, the probability (p) of detecting at least one mtDNA iBoF haplotype in a sample is

$$p=1-(1-f)^N$$

where f is the frequency of the allele and N the number of individuals sampled. Here, p was assessed over a range of haplotype frequencies observed in the inner Bay, and for different values of N.

If all the juveniles surveyed were descended from a single female, then the sampling of the Nth offspring would provide no greater power to detect a given mtDNA variant in a population than that achieved by sampling the first offspring. In other words, if the 17 juveniles had descended from 2 female parents, then the sampling of all 17 juveniles would be equivalent to sampling 2 unrelated offspring (assuming no relatedness in previous generations). To address the effect of possible relatedness (specifically, having a female parent in common), we assessed p assuming 1) that all groupings detected by *Colony* were false and that all 17 offspring were unrelated, having descended from 17 different females (N=17); 2) that all (or nearly all) full-sib groupings identified were true (as suggested by the simulation analyses) and all or nearly all half-sib groupings artificial (consistent with results from the simulation analyses), indicating that the 17 offspring descended from 13 females (N=13); and 3) that the full-sib and half-sib groupings were correct, and that the female parent was common between the two sets of full-sib groupings nested within their respective half-sib groupings (N=11). This latter mating structure (half sib parents being female) was also hypothesized in Herbinger *et al.* (2006) and later demonstrated empirically using studies of sex-linked markers (de Mestral Bezanson *et al.*, 2009), for most half-sib groups analysed in small declining iBoF populations. Thus, p was estimated for different numbers of what we call here *effective juveniles* (11, 13, and 17), or juveniles that likely inherited mtDNA from a different (unique) female parent.

RESULTS

In the fall of 2007, 19 sample tubes were submitted to DFO for genetic analyses. Two tubes were completely empty, but small (milligram) quantities of tissue were found in the remaining 17 tubes. Base determinations at the mitochondrial ND1 site were successfully made for all 17 samples submitted (Table 3). Microsatellite genotype information was obtained for 199 of the 204 single-locus genotypes involved in this 12 locus assay of 17 Atlantic salmon.

All 17 mtDNA molecules surveyed exhibited a "C" at the ND1 4199 site. In other words, iBoF mtDNA haplotypes are likely absent in the samples of 17 juveniles submitted.

Considerable microsatellite variation was observed in this collection of 17 samples analyzed, with as many as 12 to 13 different alleles observed at individual loci (Table 3). The numbers of alleles observed directly indicate that the juveniles sampled did not descend from two or three parents, but instead require an absolute minimum of seven parents, and represent a minimum of four full-sib groupings. In fact, because these loci exhibit a fixed (non-infinite) number of alleles in the Bay of Fundy populations surveyed, the numbers of alleles observed very likely reflect contributions from well in excess of 10 different parents.

Kinship analyses partitioned the 17 multilocus genotypes into 13 full-sib and 11 half-sib groupings (Table 3). Specifically, two half-sib groups were observed (4.X and 5.X, Table 3), each consisting of two full-sib groups (4.4 and 4.5, and 5.6 and 5.7, respectively). All other groupings consisted of pure full-sib groups only (each offspring having been produced by a unique male and female parent). Additionally, most half- or full-sib groups were very small, with 10 of 13 full-sib groups, and 7 of 11 half-sib groups, consisting of single members. Two full-sib groups contained two siblings, and one full-sib group three siblings. The largest half-sib group contained four members, and the sole remaining half-sib group two members.

Note also that the sample numbers (Table 3), presumably reflecting the order in which the samples were obtained and possibly geographic proximity of individual fish, are sequential (or nearly so) in the three full-sib groups of greater than 1 individual in size, except for MIS170907SOF4; this pattern has been observed several times in samples of early juveniles (unpublished data), and may reflect limited dispersal from redds or kin cohesion following dispersal.

In the analyses of 10 simulated sets of 17 unrelated individuals, no group of three or more full-sibs was detected, and only one group of two full-sibs was observed in one simulation run (simulation 3) (Table 4). Artificial half-sib groupings of two individuals, however, were observed in all simulated datasets, and ranged in number from two to seven, with the average of 3.9 per simulation. Artificial half-sib groups of three or more were much less common, and were more likely not to be observed than to be observed in any one of the 17 simulated multilocus genotypes.

When results of the actual kinship analysis are considered in the context of the simulation analyses, it would seem likely that many of the full-sib groupings of two or more individuals observed are probably real (they occur very rarely or not at all in equivalent simulated datasets of unrelated individuals), but that the half-sib groups may or may not be real. Therefore, the juveniles sampled are probably descended from 13 females (assuming all of the full sib groupings are indeed real), but may have been produced by as few as 11 females (if the half sib groupings are real) or by 17 females if all of the juveniles are unrelated. In other words, the effective number of juveniles sampled is 11, 13 or 17.

The probability of detecting an iBoF mtDNA haplotype, assuming its presence at different frequencies and given varying effective numbers of juveniles sampled, ranged from 0.7120 to 0.9999 (Table 5). When present in a population at a frequency of 0.3 or greater, the likelihood of detection is very high (>98%), whether the individuals sampled are unrelated, or given the levels of full- or half-sib structuring reported by *Colony* for the Mispec sample collection. When allele frequencies decline to approximately 0.1, such tests may fail to detect an inner Bay haplotype in about 1/4 of analyses of a similar size (17 individuals), with a similar level of family structuring observed at the full-sib level (effectively, 13 individuals). Possible further half-sib structuring would reduce the likelihood of detecting iBoF haplotypes slightly more.

In the analysis of microsatellite variation carried out here, two of the 17 Mispec samples exhibited an allele at locus *Ssa*202 247 bp in length. This allele was not observed in any of the near 2,000 Atlantic salmon, representing six iBoF populations, sampled in multiple years preceding the fall of 2000, surveyed by Jones (2001).

DISCUSSION

Since all 17 juveniles surveyed exhibited a "C" base at the mitochondrial ND1 4199 site, a straightforward conclusion of this analysis is that we find no evidence to indicate that some or all of the juveniles observed in the Mispec River in 2007 were of iBoF in origin. This lack of evidence for iBoF ancestry, though, cannot be taken as definitive evidence for a non-iBoF (wild or aquaculture) origin of these samples, as not all known or true iBoF salmon exhibit an iBoF mtDNA haplotype. However, the absence of even a single iBoF mtDNA haplotype in this collection of 17 largely unrelated juveniles (11 to 17, and most probably 13 effective juveniles) indicates that they were very probably not descended from an iBoF population that exhibited an iBoF mtDNA haplotype at a frequency of 30% or more, and most likely not from an iBoF population exhibiting an iBoF mtDNA haplotype at 10% or more. Furthermore, the presence of the *Ssa*202 microsatellite allele 247 bps in length also strongly suggests a non-iBoF origin of this group. The *Ssa*202-247 bp allele was not observed in the nearly 2,000 and 1,000 iBoF salmon obtained prior to the year 2000 by Jones (2001) or O'Reilly (unpublished data), respectively, yet was reported by Jones (2001) in the Margaree River (which drains into the Northumberland Strait) and in all drainages of the Saint John River surveyed to date, including the Serpentine, Tobique, Hammond, and Nashwaak at a maximum frequency of 4% in a sample of late-run adults obtained from the latter tributary (O'Reilly, unpublished data).

Given the geographic proximity of the large Saint John River system, which still harbours a relatively large (> 2,000 adults annually) population of Atlantic salmon (Jones *et al.*, 2006), the fact that the Mispec is the nearest river to the east of the Saint John, and that stray rates in Atlantic salmon as high as 6% have been recorded (Jonsson *et al.*, 2003), adults from this population are a very plausible source of the juveniles observed in the Mispec River in 2007. The proximity of the Mispec to Passamaquoddy and Cobscook Bay areas also suggest a possible aquaculture origin of the parents of the Mispec juveniles observed.

The kinship pattern observed here - little if any structuring at the half sib level - is markedly different from patterns observed by Herbinger *et al.* (2006) and de Mestral Bezanson *et al.* (2009) for sample collections obtained in the year 2000 from small but still persisting iBoF populations. In these two studies, juveniles sampled mostly clustered into relatively few large half-sib families comprised of many small full-sib families, with the half-sib parents reflecting a handful of returning female adults and the full-sib parents mostly small numerous mature male parr still in the system (Herbinger *et al.*, 2006; de Mestral Bezanson *et al.*, 2009). The pattern observed here is more consistent with a handful of strays or escapes ascending and spawning in an extirpated river, in the absence of large numbers of mature male parr produced by spawnings in the rivers in previous years. The mating structure reported here, and implications as to the origin of this set of samples, is also consistent with previous electrofishing surveys by Gibson *et al.* (2003) carried out in 2002 that detected few if any Atlantic salmon juveniles in any non-supported iBoF river. This pattern is also consistent with these fish being produced by a directed paired-spawning process, where single male adults are crossed once with single females. This latter scenario would involve the release of artificially produced juvenile salmon into the Mispec River.

LITERATURE CITED

Butler, K., Field, C., Herbinger, C.M., and Smith, B.R. 2004. Accuracy, efficiency and robustness of four algorithms allowing full-sibship reconstruction from DNA marker data. *Molecular Ecol.* 13: 1589–1600.

de Mestral Bezanson, L., Herbinger, C.M., and O'Reilly, P.T. 2009. Mating structure of an endangered population of wild Atlantic salmon *(Salmo salar)* as determined by sibship analysis and a novel method of gender inference. CJFAS (submitted).

DFO, 1999. Interaction between wild and farmed Atlantic salmon in the Maritime Provinces. DFO Mar. Reg. Hab. Status Rep. 99/1E.

DFO, 2006. Department of Fisheries and Oceans. Recovery Strategy for the Atlantic salmon *(Salmo salar)*, inner Bay of Fundy populations [Proposed]. *In:* Species at Risk Act Recovery Strategy Series. Ottawa: Fisheries and Oceans Canada. 62 pp.

Ducharme, L.J.A. 1969. Atlantic salmon returning for their fifth and sixth consecutive spawning trips. *J. Fish Res. Bd. Canada* 26: 1661–1664.

Gibson, A.J.F., Amiro, P.G., and Robichaud-LeBlanc, K.A. 2003. Densities of juvenile Atlantic salmon *(Salmo salar)* in inner Bay of Fundy rivers during 2000 and 2002 with reference to past abundance inferred from catch statistics and electrofishing surveys. *Can. Stock Assess. Sec. Res. Doc.* 2003/121. ii + 61p.

Hansen, M.M., and Jensen, L.G. 2005. Sibship within samples of brown trout *(Salmo trutta)* and implications for supportive breeding. *Conservation Genetics* 6: 297-305.

Herbinger, C., O'Reilly, P.T., and Verspoor, E. 2006. Unravelling pedigrees in wild endangered salmon populations using molecular genetic markers. *Molecular Ecol.* 15: 2261-2275.

Hurst, C.D., Barlett, S.E., Davidson, W.S. and Bruce, I..J. 1999. The complete *mtDNA* sequence of Atlantic salmon, *Salmo salar*. *Gene* 239: 237-242.

Jessop, B.M. 1976. Distribution and timing of tag recoveries from native and nonnative Atlantic salmon *(Salmo salar)* released into Big Salmon River, New Brunswick. *J. Fish Res. Bd. Canada* 33: 829–833.

Jones, M. 2001. Application of genetic markers to conservation biology: salmonid fish in Eastern Canada. Ph.D. thesis, Dalhouise University, Halifax, N.S.

Jones, R.A., Anderson, L., Flanagan, J.J., and Goff, T. 2006. Assessments of Atlantic salmon stocks in southern and western New Brunswick (SFA 23), an update to 2005. *Can. Sci. Advis. Sec. Res. Doc. 2006/025. ii + 82p.*

Jonsson, B., Jonsson, N., and Hansen, L.P. 2003. Atlantic salmon straying from the river Imsa. *J. Fish Biol.* 62: 641-657.

King, T.L., Eackles, M.S., and Letcher, B.H. 2005. Microsatellite DNA markers for the study of Atlantic salmon *(Salmo salar)* kinship, population structure, and mixed-fishery analyses. *Molecular Ecology Notes* 5: 130-132.

King, T.L., Spindle, A.P., Eackles, M.S., Lubinski, B.A. and Schill, W.B. 2000. Mitochondrial DNA diversity in North American and European Atlantic salmon with emphasis on the Downeast rivers of Maine. *J. Fish Biol.* 57: 614-630.

Lyamichev, V., *et al.* 1999. Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. *Nat. Biotechnol.* 17: 292–296.

Manel, S., Gaggiotti O.E., and R.S. Waples. 2005. Assignment methods: matching biological questions with appropriate techniques. *Trends Ecol. Evol.* 20: 136–142.

O'Reilly, P., Hamilton, L., McConnell, S.K., and Wright, J.M. 1996. Rapid analysis of genetic variation in Atlantic salmon (*Salmo salar*) by PCR multiplexing of tetranucleotide and dinucleotide microsatellites. *Can. J. Fish. Aquat. Sci.* 53: 2292-2298.

Paterson, S., Piertney, S.B., Knox, D., Gilbey, J., and Verspoor, E. 2004. Characterisation and PCR multiplexing of novel highly variable tetranucleotide Atlantic salmon *(Salmo salar L.)* microsatellites. *Molecular Ecology Notes* 4: 160-162.

Smith, B.R., Herbinger, C.M., and Merry, H.R. 2001. Accurate partition of individuals into full-sib families from genetic data without parental information. *Genetics* 158: 1329–1338.

Verspoor, E., O'Sullivan, M., Arnold, A.L., Knox, D. and Amiro, P. 2002. Restricted matrilineal gene flow and regional differentiation among Atlantic salmon *(Salmo salar L.)* populations within the Bay of Fundy, eastern Canada. *Heredity* 89: 465-472.

Verspoor, E., O'Sullivan, M., Arnold, A.L., Knox, D., Curry, A., Lacroix, G., and Amiro, P. 2005. The nature and distribution of genetic variation at the mitochondrial ND1 gene of the Atlantic salmon *(Salmo salar L.)* within and among rivers associated with the Bay of Fundy and the Southern Upland of Nova Scotia. FRS Research Services Internal Report No. 18/05.

Wang, J. 2004. Sibship reconstruction from genetic data with typing errors. *Genetics* 166: 1963–1979.

Table 1. Sample collection information for 19 Atlantic salmon juveniles obtained from the Mispec River and submitted to the Department of Fisheries and Oceans for genetic analyses, including length, weight, and the general and specific location at which samples were obtained*.

Vial ID (ID Submitted)	Sample ID (Laboratory ID)	Date Captured	Total Length (cm)	Fork Length (cm)	Weight (gm)	River (General Location)	UTM (Specific Location) Easting	UTM (Specific Location) Northing
1	MIS070907SOF1	07-Sep-07	13.7	12.8	20.9	Mispec River	272919	5017045
2	MIS170907SOF2	17-Sep-07	17.6	17.0	6.1	Mispec River	268408	5013757
3	MIS170907SOF3	17-Sep-07	17.0	16.2	50.1	Mispec River	268408	5013757
4	MIS170907SOF4	17-Sep-07	7.0	6.2	4.5	Mispec River	268408	5013757
5	MIS170907SOF5	17-Sep-07	7.1	6.4	3.9	Mispec River	268408	5013757
6	MIS180907SOF6	18-Sep-07	16.6	16.4	46.4	Mispec River	268793	5014043
7	MIS180907SOF7	18-Sep-07	7.0	6.3	3.7	Mispec River	268793	5014043
8	N/A+	18-Sep-07	6.3	6.3	3.1	Mispec River	268793	5014043
9	MIS180907SOF9	18-Sep-07	12.2	11.1	15.3	Mispec River	268793	5014043
10	MIS100907SOF10	10-Sep-07	6.8	6.3	2.4	Trib. 1 of Mispec R.	268560	5014081
11	N/A+	10-Sep-07	6.6	6.1	2.5	Trib. 1 of Mispec R.	268560	5014081
12	MIS260907SOF12	26-Sep-07	7.5	7.2	3.9	Trib. 1 of Mispec R.	268557	5013988
13	MIS260907SOF13	26-Sep-07	6.5	6.1	3.1	Trib. 1 of Mispec R.	266557	5013988
14	MIS260907SOF14	26-Sep-07	15.4	14.9	34.2	Trib. 1 of Mispec R.	268557	5013988
15	MIS260907SOF15	26-Sep-07	4.8	4.4	1.7	Mispec River	272919	5017045
16	MIS260907SOF16	26-Sep-07	5.3	4.9	1.0	Mispec River	272919	5017045
17	MIS260907SOF17	26-Sep-07	5.5	5.0	1.2	Mispec River	272919	5017045
18	MIS260907SOF18	26-Sep-07	5.4	4.9	1.5	Mispec River	272919	5017045
19	MIS260907SOF19	26-Sep-07	5.0	4.7	1.2	Mispec River	272919	5017045

* Samples were submitted by Lee Jamieson, Jacques Whitford Consultants, November 7, 2007.
N/A+ Tissue amounts, and therefore expected DNA quantity, were too small to permit genetic analyses

Table 2. Marker type employed (nuclear versus mitochondrial), mode of inheritance, number of alleles observed, and locus references providing primer sequences and annealing temperatures for microsatellite loci.

Locus	Type/Inheritance Pattern	Number of Alleles	Reference (s)
58	Nuclear/bi-parental	13	King *et al.*, 2005
144	Nuclear/bi-parental	10	King *et al.*, 2005
171	Nuclear/bi-parental	11	O'Reilly *et al.*, 1996
197	Nuclear/bi-parental	8	O'Reilly *et al.*, 1996
202	Nuclear/bi-parental	9	O'Reilly *et al.*, 1996
486	Nuclear/bi-parental	6	King *et al.*, 2005
1605	Nuclear/bi-parental	6	Paterson *et al.*, 2004
2201	Nuclear/bi-parental	13	Paterson *et al.*, 2004
2210	Nuclear/bi-parental	6	Paterson *et al.*, 2004
2215	Nuclear/bi-parental	7	Paterson *et al.*, 2004
2216	Nuclear/bi-parental	12	Paterson *et al.*, 2004
1G7	Nuclear/bi-parental	6	Paterson *et al.*, 2004
ND1,Site 4199	Mitochondrial /Maternal	1	Hurst *et al.*,1999; Verspoor *et al.*, 2002.

Table 3. Microsatellite and mitochondiral DNA (mtDNA) Single Nucleotide Polymorphism (SNP) data, and possible first-order relatedness as inferred from Kinship analyses, for 17 Atlantic salmon samples analyzed.

	Microsatellite Locus																									
Sample ID	SSaD58	SSaD58	SSaD144	SSaD144	Ssa171	Ssa171	Ssa197	Ssa197	Ssa202	Ssa202	SSaD486	SSaD486	SSsp1605	SSsp1605	SSsp2201	SSsp2201	SSsp2210	SSsp2210	SSsp2215	SSsp2215	SSsp2216	SSsp2216	SSsp1G7	SSsp1G7	mt DNA	KIN GRP
MIS070907SOF1	336	368	162	186	233	239	175	187	291	295	190	190	236	256	287	351	124	136	162	162	201	209	198	218	C	1.1
MIS260907SOF14	340	356	214	214	231	255	167	195	279	299	174	190	232	232	279	355	120	124	166	166	205	397	182	206	C	2.2
MIS260907SOF16	320	326	178	202	245	249	191	195	247	283	178	178	232	232	243	295	112	120	154	162	233	265	198	202	C	3.3
MIS260907SOF18	326	340	242	298	235	235	163	195	247	283	174	174	232	232	243	335	120	124	154	170	229	261	198	202	C	3.3
MIS170907SOF4	336	368	190	198	229	243	171	171	291	307	178	194	228	236	295	331	112	124	166	190	249	265	214	214	C	4.4
MIS260907SOF12	352	372	162	186	249	257	163	171	299	307	178	194	228	228	311	331	112	128	166	190	209	265	178	182	C	4.4
MIS260907SOF13	352	372	162	198	229	249	163	167	299	307	178	210	228	236	311	331	112	124	166	190	249	265	178	182	C	4.4
MIS180907SOF7					229	235	163	195	307	311	174	194	228	228	295	355	112	136	154	190	261	265	206	214	C	4.5
MIS260907SOF15	332	340	202	298	235	257	179	179	303	311	178	198	232	264	295	335	124	128	150	170	241	261	182	198	C	5.6
MIS260907SOF17	320	320	198	202	233	257	167	187	299	311	174	198	228	232	295	335	112	112	162	170	233	261	182	198	C	5.7
MIS170907SOF2	368	388			249	249	163	179	299	307	194	210	236	236	295	315	120	128	162	170	205	229	174	182	C	6.8
MIS170907SOF3	368	388	162	198	249	263	167	171	295	299	194	194	232	236	299	311	112	128	166	170	205	229	182	214	C	6.8
MIS170907SOF5	364	372	198	230			163	171	299	311	194	210	228	232	311	355	132	136	162	166	201	249	182	182	C	7.9
MIS260907SOF19	340	356			229	243	167	179	307	311	174	174	236	252	243	335	112	124	158	170	257	261	174	182	C	8.10
MIS100907SOF10	336	340	162	198	243	249	163	171	283	311	178	194	232	236	299	331	112	124	162	170	201	257	166	182	C	9.11
MIS180907SOF6	312	388	162	186	249	255	171	179	283	299	174	194	232	232	295	363	124	136	162	190	209	249	174	182	C	10.12
MIS180907SOF9	320	380	198	298	229	243	167	175	295	295	174	178	232	236	291	311	112	112	158	170	253	265	182	198	C	11.13

KIN GRP (Kinship group) designations: The number preceding the decimal reflects the half-sib family, and the number following the decimal the full-sib family nested within.

Table 4. Distribution of 17 multilocus genotypes into full- and half-sib groupings for each of 10 simulated datasets.

Simulation Number	Number of Full-sib Groups of Two	Number of Full-sib Groups of Three or More	Number of Half-sib Groups of Two	Number of Half-sib Groups of Three or More
1	0	0	4	0
2	0	0	2	0
3	1	0	3	1
4	0	0	7	0
5	0	0	5	0
6	0	0	4	1
7	0	0	3	0
8	0	0	3	1
9	0	0	5	0
10	0	0	3	3
Average	0.1	0	3.9	0.6

Table 5. Likelihood of detecting inner Bay of Fundy mitochondrial DNA types given variable population allele frequencies and different effective numbers of juveniles sampled.

Haplotype Frequency*	Effective Number of Juveniles Sampled**	Kin Structuring Associated with Numbers of Female Parents Sampled	Probability of Detecting at Least one iBoF mtDNA Haplotype
0.107	17	No relatedness	0.8540
(Big Salmon River)	13	Observed full-sib groupings	0.7703
(P. O'Reilly, unpub. data)	11	Observed half-sib groupings	0.7120
0.313	17	No relatedness	0.9983
(Stewiacke R.)	13	Observed full-sib groupings	0.9924
(Verspoor *et al.*, 2002)	11	Observed half-sib groupings	0.9839
0.500	17	No relatedness	0.9999
(Folly R.)	13	Observed full-sib groupings	0.9998
(Verspoor *et al.*, 2002)	11	Observed half-sib groupings	0.9995
0.179	17	No relatedness	0.9643
(inner Bay of Fundy)	13	Observed full-sib groupings	0.9218
(Verspoor *et al.*, 2002)	11	Observed half-sib groupings	0.8842

Probability (p) of detecting at least one mtDNA haplotype
$p = 1-(1-f)^N$, given a frequency of f and a sample size of N.

* Population allele frequencies for the inner Bay mtDNA haplotype used above reflect the range observed by Verspoor *et al.* (2002), where iBoF mtDNA haplotypes were present.

** The effective number of juveniles are the numbers of juveniles that likely reflect the mtDNA haplotype of a different and unique female parent, under a given scenario (described more fully in the Methods section).

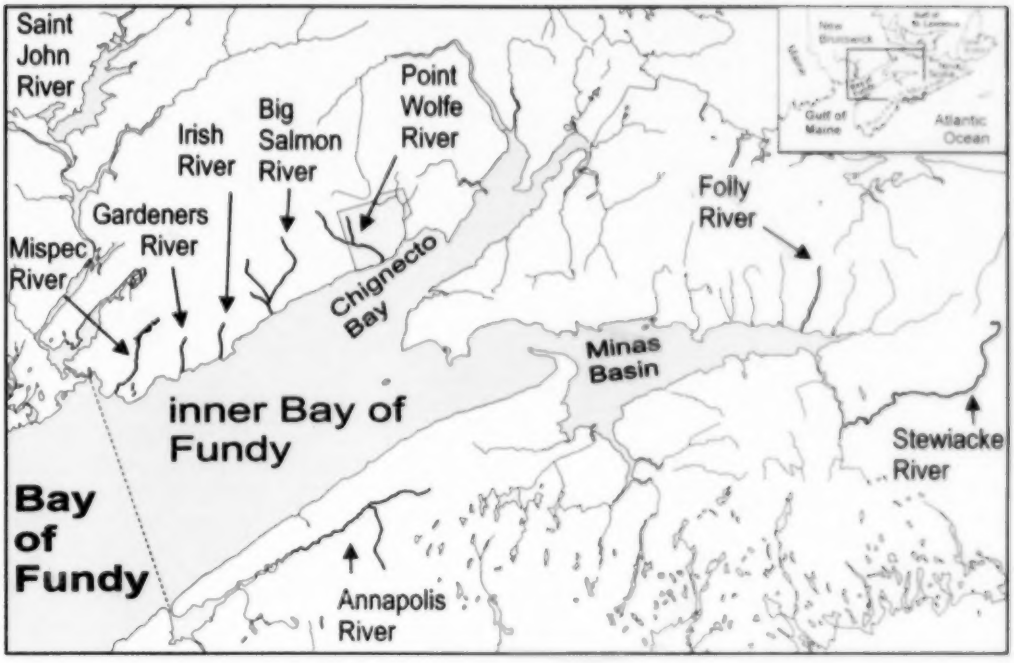

Figure 1. Geographic locations of the Mispec and other reference rivers in the Bay of Fundy area.